



CHRIST

(DEEMED TO BE UNIVERSITY)

PUNE LAVASA CAMPUS

The Hub of Analytics

Internship Report

by

FATHIMATHUL SUSNA SHOUKKATHALI

(20112013)

Under the guidance of

Prof. Shynu Philip

&

Mr. Liqzan Manna

**A Internship Report submitted in partial fulfilment of the
requirements for the award of Degree of Bachelor of Science**

(Data Science)

of

CHRIST (Deemed to be University),

Pune Lavasa Campus

Academic Year 2021-2022



CHRIST

(DEEMED TO BE UNIVERSITY)

PUNE LAVASA CAMPUS

The Hub of Analytics

CERTIFICATE

*This is to certify that the report titled **Internship Report** is a bonafide record of work done by **Fathimathul Susna Shoukkathali (20112013)** of CHRIST (Deemed to be University), Pune Lavasa Campus, in partial fulfillment of the requirements of III Semester BSc (Data Science) during the year 2021-2022.*

Head of the Department

Dr. Samiksha Shukla

Guide

Prof. Shynu Philip

Internship Certificate

**smartknower**
STAY AHEAD OF THE CURVE

E-Cell
IIT Hyderabad

CERTIFICATE OF COURSE COMPLETION

THIS CERTIFICATE IS PROUDLY PRESENTED TO

Fathimathul Susna S Ali

Has successfully completed course on Machine Learning
from the period of 1/6/2021 to 31/7/2021.
During his/her training we found the student was dedicated and hard working.

13-08-2021
DATE



P. Naveen
ACADEMIC HEAD

Certificate ID : 2289624167

**smartknower**
STAY AHEAD OF THE CURVE

Best
AWARD

CERTIFICATE OF INTERNSHIP

THIS CERTIFICATE IS PROUDLY PRESENTED TO:

Fathimathul Susna S Ali

Has successfully completed Internship On Machine Learning
from 1/6/2021 to 31/7/2021. During his/her internship,
the student was found to be dedicated, hardworking and intelligent

20-08-2021
DATE

P. Naveen
Academic Head

Certificate ID: 1937317247

CONTENTS

| S. No | Title of Content | Page No |
|--------------|---|----------------|
| | Acknowledgement | I |
| 1. | Introduction | 6 |
| 2. | Description of the Internship 2.1 Organization Details 2.2 Internship Activities | 7-8 |
| 3. | Reflection on Internship | 9 |
| 4. | Conclusion | 10 |
| | References | 11 |
| | Appendix | 12-14 |

ACKNOWLEDGEMENT

The internship opportunity I had with Smartknower was a great experience for learning and enhancing my skills. Therefore, I consider myself lucky individual as I was provided with the opportunity to be a part of this course. I am also grateful to meet many great people who led me through this internship course period.

First and foremost, I would like to acknowledge my sincere gratitude towards “The Almighty” for granting me His protection, opportunity, and strength to complete this internship.

Secondly, I express my deep and sincere thanks to **Fr. Jossy P George** Director of **Christ (Deemed to be University)**, **Pune Lavasa** for arranging an Internship programme, also I would like to express my gratitude to **Dr. Samiksha Shukla**, HOD of Data Science Department, Our Co-ordinator **Prof. Pramila R M** and **Prof. Shynu Philip** who supported me to prepare and submit this report in a due time.

I express my heartfelt gratitude and sincere thanks to internship mentor and guide. **Mr Liqzan Manna**, despite of his busy schedule, took time to hear out, guide, for providing me with the most valuable guidance and affable treatment given to me at every stage to boost my morale and extend during this internship.

I am extremely grateful to my department staff members and friends who helped me in the successful completion of this internship.

1.Introduction

The internship programme was carried out at the Smartknower, Bangalore under the guidance of Mr. Liqzan Manna who offered us demanding and immersive training along with simulated real-life technical exposure through live projects. As mentioned, I had opted Machine Learning Course which covered the concepts and basics of python Machine Learning packages with their algorithms, understand the importance of Machine Learning and how it is applicable in our daily lives, implementing ML algorithm on datasets to recognise and analyse programmes. They have provided us with a glimpse idea of Natural Language Processing (NLP). Once the concepts were covered, I was provided with two projects.

2.Description of the Internship

2.1 Organization Details

Smartknower is an organization founded in the year 2020 located in Bangalore, dedicated to the field of educational technology which promotes E-Learning that makes accessible for everyone at any part of the world.

Their vision and mission are to enhance students' life by specializing in modern technology and entrepreneurial domains that helps to benefit from these programs to become a productive and proactive part of the co-operate space. Rigorous and immersive training, as well as real-life technical exposure through live projects, guides the candidates to develop their comprehensive capabilities.

Weekly Mentoring sessions with industry experts' assistance clears doubts and explains the topics in an understandable manner. The ground-level employ technique approach in training students in the concepts pertaining to the domain. The company also provides the candidate with an interconnected ecosystem that can be accessed from anywhere which certainly helped in this pandemic situation.

2.2 Internship Activities

This course consisted of the key concepts and techniques of python Machine Learning packages which are used to design and learn from basics to advanced.

They started from the very basics of python where they started by giving us the background and the importance of the language. Then they gave us the introduction to the fundamentals and some of the packages which could be used for Machine Learning algorithms. Learned the importance of Machine Learning and how we use it in our daily tasks. The types of Machine learning and its significance, taught us with the difference of Supervised Learning and Unsupervised Learning where we used Linear algorithm and Logistic regression algorithm and introduced us to Support Vector Machine (SVM) under supervised learning for analysis. In case of unsupervised learning, we used Apriori algorithm, taught us the basics and introduced us to KNN and Clustering. We used Scikit-learn library which is a simple and effective tool that helps for predictive analysis, it also consists of unsupervised and supervised learning algorithms which we used for our analysis and prediction. Finally with the basics of WebApp for Deployment, taught us how to write ML algorithms and the codes that would make the programs more intellectual. They introduced Natural Language Programming (NLP), which facilitates the connection between machines and human beings.

During the course, they provided us with two projects where one was to perform any classification algorithm and compare the accuracy on the iris dataset. The other project consists of a National Institute of Diabetes and Digestive and Kidney Diseases dataset which is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. They asked us to build a machine learning model and Webapp (streamlit) to accurately predict whether or not the patients in the dataset have diabetes or not.

3. Reflection on Internship

In the past few years, especially thanking the recent achievements in the field of Machine Learning has drawn a lot of attention around the world which made me curious to know more about this field. The main driving factor is associated with the fact that it bids a unified framework for presenting intelligent decision-making into many fields.

With the help of this internship, I have got introduced to the field of Machine Learning, specifically focusing on the core concepts of the supervised and unsupervised learning. The ability to design and implement various machine learning algorithm in a range of dataset related to real- world applications.

I have gained great knowledge of Machine Learning and the application of its algorithm on the dataset. Understood the importance of the Machine Learning in our daily life and its necessity. The main learning outcome of this internship is that now I can create basic Machine Learning models, I can analyse and predict the accuracy of the dataset provided. With the live projects provided by the organization, I was able to measure the understanding, demonstrate the personal abilities and reflect on learning. Learning these helps us to solve new tasks, related to the preceding encountered tasks more effectually.

4.Conclusion

As expressed, this internship has been an excellent and useful experience. I have gained new knowledge, skills and achieved my learning goal. Most importantly I was able to gain valuable insights from professionals and their guidance was an important factor that led this internship smoothly.

This internship pointed us to work with our strengths and weaknesses that helped us to define the knowledge and skills that have to be improved in the future. Moreover, the internship helps to learn independently, be considerate, patient and the ability to solve the problems.

Thus, after the two months of hard work, I was confident enough to complete my projects at my own by satisfying the requirements of the mentor and the organization. The visions we gained throughout this internship are adequate to contribute towards our future endeavours.

References

<https://www.analyticsvidhya.com/blog/2021/06/build-web-app-instantly-for-machine-learning-using-streamlit/> - Streamlit for creating webapp for Machine Learning.

<https://www.smartkowner.com/> - Smartkowner.

Appendix

Class attended online:

The screenshot shows a Zoom window with a hand-drawn table and formula. The table has 6 columns: x , y , $x - \bar{x}$, $y - \bar{y}$, $(x - \bar{x})^2$, and $(x - \bar{x})(y - \bar{y})$. The rows represent data points, with the last row showing the sums for each column. Below the table, the formula for the slope b_1 is written as
$$\text{slope} = b_1 = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

| x | y | $x - \bar{x}$ | $y - \bar{y}$ | $(x - \bar{x})^2$ | $(x - \bar{x})(y - \bar{y})$ |
|-----|-----|---------------|---------------|-------------------|------------------------------|
| 1 | 2 | -2 | -2 | 4 | 4 |
| 2 | 4 | -1 | 0 | 1 | 0 |
| 3 | 5 | 0 | 1 | 0 | 0 |
| 4 | 4 | 1 | 0 | 1 | 0 |
| 5 | 5 | 2 | 1 | 4 | 2 |
| 3 | 4 | | | 10 | 6 |

Below the table, the formula for the slope is written:

$$\text{slope} = b_1 = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2}$$

First Project Screenshot:

The screenshot shows a Jupyter Notebook interface with the following code and output:

```
In [21]: #Gives training data to the testing algorithm
pred_model.predict(x_test)
from sklearn.metrics import classification_report, confusion_matrix
print(confusion_matrix(y_test, pred))
#Checking the accuracy of the data
accuracy_score(y_test, pred)
```

```
Out[21]: 1.0
```

Classification Report

```
In [22]: from sklearn.metrics import classification_report
classreport=classification_report(y_test,y_pred)
print(classreport)
```

| | precision | recall | f1-score | support |
|-----------------|-----------|--------|----------|---------|
| Iris-setosa | 1.00 | 1.00 | 1.00 | 11 |
| Iris-versicolor | 1.00 | 1.00 | 1.00 | 13 |
| Iris-virginica | 1.00 | 1.00 | 1.00 | 6 |
| accuracy | | | 1.00 | 30 |
| macro avg | 1.00 | 1.00 | 1.00 | 30 |
| weighted avg | 1.00 | 1.00 | 1.00 | 30 |

Here I have converted my. ipynb to .py and then runned it so I was able to create the App Screenshot.

```

Anaconda Prompt (anaconda3) - streamlit run MajorProject.py
(base) C:\Users\Fathi\2nd Semester>ipynb-py-convert MajorProject.ipynb MajorProject.py
(base) C:\Users\Fathi\2nd Semester>streamlit run MajorProject.py

You can now view your Streamlit app in your browser.

Local URL: http://localhost:8501
Network URL: http://172.31.243.133:8501

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
Pregnancies      768 non-null int64
Glucose          768 non-null int64
BloodPressure    768 non-null int64
SkinThickness    768 non-null int64
Insulin          768 non-null int64
BMI              768 non-null float64
DiabetesPedigreeFunction 768 non-null float64
Age              768 non-null int64
Outcome          768 non-null int64
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
2021-08-13 11:57:21.284 NumExpr defaulting to 8 threads.
c:\users\fathi\anaconda3\lib\site-packages\pandas\plotting\_matplotlib\tools.py:307: MatplotlibDeprecationWarning:
The rowNum attribute was deprecated in Matplotlib 3.2 and will be removed two minor releases later. Use ax.get_subplotspec().rowspan.start instead.
  layout[ax.rowNum, ax.colNum] = ax.get_visible()
c:\users\fathi\anaconda3\lib\site-packages\pandas\plotting\_matplotlib\tools.py:307: MatplotlibDeprecationWarning:
The colNum attribute was deprecated in Matplotlib 3.2 and will be removed two minor releases later. Use ax.get_subplotspec().colspan.start instead.
  layout[ax.rowNum, ax.colNum] = ax.get_visible()
c:\users\fathi\anaconda3\lib\site-packages\pandas\plotting\_matplotlib\tools.py:313: MatplotlibDeprecationWarning:
The rowNum attribute was deprecated in Matplotlib 3.2 and will be removed two minor releases later. Use ax.get_subplotspec().rowspan.start instead.
  if not layout[ax.rowNum + 1, ax.colNum]:
c:\users\fathi\anaconda3\lib\site-packages\pandas\plotting\_matplotlib\tools.py:313: MatplotlibDeprecationWarning:
The colNum attribute was deprecated in Matplotlib 3.2 and will be removed two minor releases later. Use ax.get_subplotspec().colspan.start instead.
  if not layout[ax.rowNum + 1, ax.colNum]:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
Pregnancies      768 non-null int64
Glucose          768 non-null int64
BloodPressure    768 non-null int64

```

The App I created for my second Project screenshots.

Fathimathul Susna ShoukkathAli

Diabetes Prediction

Select Form

Form 1

☐ Hide

Triceps skin fold thickness (mm):

35.00

2-Hour serum insulin (mu U/ml):

79.80

Body mass index (weight in kg/(height in m)^2):

33.58

Diabetes Pedigree Function:

0.63

Age:

50.00

Predict

| | Pregnancies | Glucose | BloodPressure | SkinThickness | Insulin | BMI | Di |
|---|-------------|----------|---------------|---------------|---------|---------|----|
| 0 | 6 | 148.0000 | 72.0000 | 35.0000 | 79.7995 | 33.6000 | |
| 1 | 1 | 85.0000 | 66.0000 | 29.0000 | 79.7995 | 26.6000 | |

